Quantifying Prediction Fidelity in Ocean CirculationModels

Mohamed Iskandarani
Rosenstiel School of Marine and Atmoshperic Science
University of Miami
4600 Rickenbacker Causeway, Miami, FL 33149
phone: (305) 421-4045 fax: (305) 421-4696 email:miskandarani@rsmas.miami.edu

Omar Knio
Department of Mechanical Engineering and Materials Science
Duke University
144 Hudson Hall, Box 90300, Durham, NC 27708
phone: (919) 660 5344 fax: (919) 660-8963 email:omar.knio@duke.edu

Ashwanth Srinivasan
Tendral LLC
77 Crandon Blvd, Key Biscayne, FL, 33149-1403
phone: (786)213-5201 email:asrinivasan@rsmas.miami.edu

William C. Thacker
Cooperative Institute for Marine and Atmospheric Studies University of Miami
4600 Rickenbacker Causeway, Miami, FL 33149
phone: (305) 361-4323 fax: (305) 421-4696 email:carlisle.thacker@noaa.gov

Award Number: N00014-101-0498
http://www.onr.navy.mil/sci_tech/32/reports/annual/

LONG-TERM GOALS

The development, validation and application of robust uncertainty quantification methods to ocean modeling, forecasting, and parameter estimation.

OBJECTIVES

This project explores the use of Polynomial Chaos (PC) expansions for improving our understanding of uncertainties in Ocean General Circulation Models (OGCM). Reliable ocean forecasts require an objective, practical and accurate methodology to assess the inherent uncertainties associated with the model *and* data used to produce these forecasts. OGCMs uncertainties stem from several sources that include: physical approximation of the governing equations; discretization and modeling errors; an incomplete set of sparse (and often noisy) observations to constrain the initial and boundary conditions of the model; and uncertainties in surface momentum and buoyancy fluxes.

Report Documentation Page	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE **30 SEP 2013**	2. REPORT TYPE	3. DATES COVERED **00-00-2013 to 00-00-2013**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Quantifying Prediction Fidelity in Ocean CirculationModels	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Miami,Rosenstiel School of Marine and Atmoshperic Science (RSMAS),4600 Rickenbacker Causeway,Miami,FL,33149**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**12**	

Our objective is the development of an uncertainty quantification methodology that is efficient in representing the solution's dependence on the stochastic data, that is robust even when the solution depends discontinuously on the stochastic inputs, that can handle non-linear processes, that propagates the full probability density functions without apriori assumption of Gaussianity, and that can be applied adaptively to probe regions of steep variations and/or bifurcation in a high-dimensional parametric space. In addition we are interested in developing utilities for decision support analysis; specifically, we plan to demonstrate how PC representations can be used effectively to determine the non-linear sensitivity of the solution to particular components of the random data, identify dominant contributors to solution uncertainty, as well as guide and prioritize the gathering of additional data through experiments or field observations.

APPROACH

Our approach to uncertainty quantification (UQ) relies on PC expansions (Le Maître and Knio (2010)) to investigate uncertainties in simulating the oceanic circulation. We have opted to use the HYbrid Coordinate Ocean Model (HYCOM) as our simulation engine because it has been developed as the next generation model for the US Navy and has been adopted by NOAAs National Center for Environmental Prediction. HYCOM is equipped with a suite of sequential assimilation schemes that will be used to investigate how UQ may be beneficial to data assimilation. Details about the model, its validation, and sample applications can be found in Bleck (2002); Halliwell (2004); Chassignet et al. (2003, 2006) as well as by visiting http://www.hycom.org. Below we present the main ideas of the PC expansion before we summarize our efforts since the last annual report of Sep 2012.

PC expansions express the dependency of the solution on the uncertain parameter as a series of the form: $u(\boldsymbol{x},t,\boldsymbol{\xi}) = \sum_{k=0}^{P} \hat{u}_k(\boldsymbol{x},t)\Psi_k(\boldsymbol{\xi})$, where $u(\boldsymbol{x},t,\boldsymbol{\xi})$ is a model solution that depends on space $\boldsymbol{x}$, time t and the uncertain parameters $\boldsymbol{\xi}$; $\Psi_k(\boldsymbol{\xi})$ is a suitably chosen orthogonal basis; and $\hat{u}_k(\boldsymbol{x},t)$ are the expansion coefficients. Here, u can represent a variable expressed directly in the model such as sea surface temperature or velocity at a specified point, or a derived quantity such as the mean surface cooling under a hurricane track. In the jargon of UQ u is referred to as a Quantity of Interest or an observable.

The choice of basis function is dictated primarily by the probability density function of the uncertain input data, $p(\boldsymbol{\xi})$, which enters all aspects of the UQ computations. These can be done much more efficiently if the basis vectors are orthonormal with respect to $p(\boldsymbol{\xi})$. Hence the basis functions are Legendre, Hermite, or Laguerre polynomials when the input uncertainty is described by uniform, Gaussian, or Gamma distributions, respectively.

The computation of the stochastic modes is best achieved by the so-called Non-Intrusive Spectral Projection (NISP) method since we would like to avoid modifying the original OGCM code. Taking advantage of the orthonormality of the basis, NISP works by projecting the solution u on the basis function Ψ_k via inner products, and by replacing the integrals with quadrature formula. The coefficients can then be computed simply by running the model at specified values of the uncertain parameters, storing the desired observable, and post-processing via a simple matrix-vector multiplication. No modification to the OGCM need be performed.

The investigative team at Duke University consisted of Dr. Omar M. Knio, and his post-doctoral associates, Drs. Alen Alexanderian and Ihab Sraj, and graduate student, Justin Winokur; they have concentrated on advancing the technical and theoretical aspects of the Uncertainty Quantification

efforts. Drs. Mohamed Iskandarani (lead PI), Ashwanth Srinivasan and William C. Thacker (University of Miami) have focused on formulating the oceanographic uncertainty problems, the modification to the HYCOM code and its actual execution, and the preparation of the necessary data to carry out the research agenda. Dr. Matthieu Le Henaff assisted us with the Gulf of Mexico configuration, and we have held discussions with Dr. François Counillon as to the applicability of PCs in Ensemble Kalman Filters based data assimilation. In FY12, we collaborated with Dr. Shuyi Chen on the inverse modelling problem discussed below; her research group produced the high-resolution space time atmospheric fields needed to force HYCOM during typhoon Fanapi. In FY12, we also collaborated with Dr. Youssef Marzouk's team at MIT on the development and implementation of sparse, adaptive, pseudo-spectral quadratures. Two graduate students, Rafael Goncalves and Shitao Wang, are now applying the uncertainty quantifications tools to oil spill modelling problems.

WORK COMPLETED

Variational Drag-Parameter Estimation

We have continued working on estimating the drag parameters from AXBT data collected during Typhoon Fanapi; but we have developed an alternative solution strategy for the inverse problem which reuses the PC surrogate within the context of a variational formulation. Our previous approach, Sraj et al. (2013), used a Bayesian inference framework, and relied on Markov Chain Monte-Carlo (MCMC) to construct the parameters full posterior distributions. The approach's efficiency hinged on a faithful Polynomial Chaos (PC) surrogate to circumvent the large computational cost associated with the MCMC sampling (each sample was the equivalent of a forward HYCOM[1] run and 10^6 samples were used).

The full construction of the posterior may be unnecessary if one is merely interested in the mode of the posterior distribution and the uncertainty around it. The posteriors' center and spread are generally sufficient to gauge the value and uncertainty of the optimal parameters, and they can be obtained straightforwardly using the PC series. Furthermore, a full distribution could be constructed if the latter is assumed to be approximately Gaussian. The advantages of this alternative strategy is that the MCMC costs and complications can be avoided, and the PC series can be used to compute the cost function gradients without any need for an adjoint code (only forward model runs are needed to build the surrogate).

To this end, the inverse problem in Sraj et al. (2013) is first recast as the minimization of the log-likelihood cost function penalizing the misfit between predictions and observations; an optimization algorithm is then applied to obtain the solution. A major hurdle is in computing the gradients needed during optimization, and which usually necessitates the tedious development and application of an adjoint code. Here we completely bypass this hurdle by reusing the surrogate developed in Sraj et al. (2013) to compute the necessary gradients. The PC series also delivers the useful but hard to compute cost function's Hessian at very little extra computational cost. The Hessian can be used to enhance the robustness and performance of the minimization algorithm, and to provide an estimate of the spread of the posterior distribution around the optimal values. Once the surrogate is available, the minimization algorithm can proceed without any additional model run. The methodology is applicable to a wide range of atmospheric and oceanic models, and has been illustrated here for HYCOM, a full three-dimensional and complex Ocean General Circulation model. This work has been summarized in a

[1] HYbrid Coordinate Ocean Model

manuscript that is currently under review (Sraj et al. (2013, in review)).

Initial Condition Uncertainty
We have also been working on using PC series to quantify the initial condition uncertainties in HYCOM oceanic forecasts in the Gulf of Mexico. The main challenge here concerns addressing uncertainty in a continuous field using a handful of uncertain parameters only. We have adopted Empirical Orthogonal Functions (EOFs) as our main tool to effect this “compression” of uncertain variables. These EOFs are obtained from Gulf of Mexico Hycom simulations, which, can be used as characteristic modes of variability in the Gulf of Mexico. These modes are multiplied by a stochastic amplitude and added to a control run; the stochastic amplitude are the uncertain stochastic parameters. We have experimented with EOFs from different HYCOM simulation, and we have finally settled on using a 14-day simulation to extract relevant EOFs. The uncertainty in these short runs were deemed more likely to represent ”uncertainty of the day” rather those stemming from year-long runs. Furthermore, the interest in using PC in data assimilation favors pursuing perturbation that are localized in space-time rather than ones reflecting basin-wide or seasonal dynamics. Figure 2 shows the first 2 SSH modes extracted from a 14-day simulation. The first mode can be identified with the uncertainty in the strength of a frontal eddy, whose interaction with the Loop Current is thought to play an important role in eddy shedding. This picture is corroborated in figure 3 where the stronger frontal eddy leads to an early shedding of the Loop Current. We have used a PC series with Legendre polynomials for basis function with a maximum degree of 6. Gauss-Legendre quadrature of order 6 was used to effect the projection requiring an ensemble of 49 members.

Uncertainty in Oil Spill Modelling
The GoM simulations are being readied for application in two separate but related areas. The first concerns the application of the PC series to data assimilation problems. The second concerns the quantification of uncertatinties in oil spill modelling. These efforts are in their early stages but we hope to make rapid progress soon, particularly with the help of the graduate students funded by the Gulf of Mexico Research Initiative.

The work associated with this project has been publicized at several conferences, workshops, seminars and invited talks, including:

- “Application of Polynomial Chaos Methods to Ocean Modeling” 2013 SIAM Annual Meeting, San Diego CA, Jul 8–12 2013.
- “Quantifying Initial Conditions Uncertainties in a Gulf of Mexico HYCOM Ocean Forecast” CARTHE Spring Meeting Miami, Florida 29-31 May 2013
- “Combining HYCOM, AXBTs and Polynomial Chaos Methods to Estimate Wind Drag Parameters”, Layered Ocean Models Workshop, U. Michigan at Ann Arbor, May 21–23 2013.
- “Quantifying Initial Conditions Uncertainties in a Gulf of Mexico HYCOM Forecast” 2013 Gulf of Mexico Oil Spill & Ecosystem Science Conference, 21–23 Jan 2013.
- “Bayesian Inference of Wind Drag Parameters Using ITOP Data and Polynomial Chaos Methods”, AMP-CSTAMP seminar, Dec 10 2012.
- “Bayesian inference of wind drag parameters at high wind speeds using a polynomial chaos surrogate”, International Conference on Ensemble Methods in Geophysical Sciences Toulouse, France, 12-16 November 2012 (poster).
- “Data assimilation using a polynomial chaos based ensemble”, International Conference on Ensemble Methods in Geophysical Sciences Toulouse, France, 12-16 November 2012.
- “Inverse Modeling and Sensitivity Analysis in Ocean Models using Polynomial Chaos Expansions”

Applied Mathematics and Computational Science Seminar, King Abdallah University of Technology, KSA, Oct 2012.

RESULTS

Parameter Estimation

The parameters estimated from the variational solution are listed in table 1. The optimal values of the multiplicative drag factor α and saturation wind speed V_{max} are in agreement with the *Maximum-A-Posteriori* (MAP) values obtained using the MCMC approach (Sraj et al., 2013). The apparent disagreement in the after-saturation slope m is inconsequential as the MCMC analysis showed the AXBT data to be uninformative with regard to m. The optimal hyper-parameters measuring observational errors σ_d^2 are listed in Table 1; the comparison with the MAP values shows a perfect agreement between the results of the MCMC and variational approaches.

Figure 1 shows the posteriors Gaussian fit using the variational means and spreads, and compares them to the MCMC posteriors. Those MCMC posteriors that are approximately Gaussian are in good agreements with the variational results (α and hyper-parameter posteriors), whereas significant disagreement occurs when the posteriors is far from a Gaussian shape ($V_{\max}$ for example exhibits appreciable skewness).

The variational PC approach offers an attractive solution to parameter identification problems when the following issues are relevant: the problem requires a detailed exploration of a relatively low-dimensional parameter space; an adjoint model is not available for the complex forward model; the optimization solution requires access to the Hessian and/or to a global view of the cost function (e.g. to identify local minima); and more detailed information about the posterior distribution is required than just its center (e.g. spread). Our exploration to date has shown the great utility of approximating the complex model with a series which can be later mined for either statistical inference Sraj et al. (2013) or optimial control Sraj et al. (2013, in review).

Initial Condition Uncertainty in Gulf of Mexico

Maps of standard deviations for Sea Surface Height (SSH) are shown in figure 4 and indicate that most of the uncertainty is concentrated in the Loop Current region and is directly related to the strength of the frontal eddy. The approximation error committed in replacing the model with a series is shown in figure 5 for SSH. These error measures are essential indicator for the reliability of the series. Here, the increase in error with time points to a gradual deterioration of the series so that by day 60 the maximum SSH error is almost 38% of the maximum standard deviation. This is indicative that the series' approximation property have slowly eroded and that there is additional uncertainty that needs to be accounted for. Improvements require a longer series with increased polynomial degree and additional sampling of the response surface. We are planning to pursue these line of thought but switching from a Gaussian quadrature type to an adaptive pseudo-spectral quadrature Winokur et al. (2013). In the mean time we are preparing a manuscript describing the lessons learned from this initial ensemble.

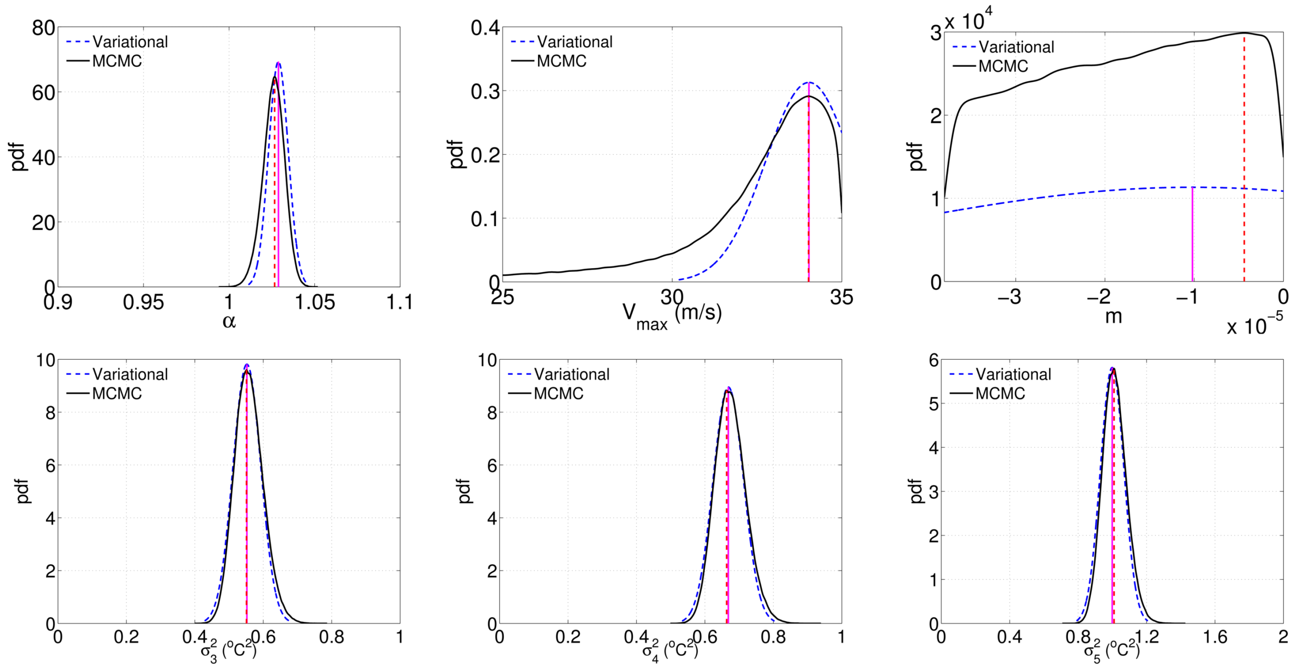

Figure 1: Posterior probability distributions for (top) drag parameters and (bottom) variances σ_d^2 at selected days using variational method (blue curves) and MCMC from Sraj et al. (2013) (black curves). The vertical lines correspond to the MAP values determined using MCMC and optimal parameters using the variational method.

Method	Variational		MCMC	
Parameter	Optimal	Spread	MAP	Spread
α	1.0289	0.0058	1.0267	0.0064
V_{max}	34.0314	1.2754	34.0190	2.4354
m	-1.0195×10^{-5}	3.5214×10^{-5}	-0.4394×10^{-5}	1.0824×10^{-5}
σ_1^2	0.6554	0.0637	0.6536	0.0655
σ_2^2	0.5712	0.0435	0.5699	0.0445
σ_3^2	0.5522	0.0407	0.5578	0.0418
σ_4^2	0.6684	0.0446	0.6742	0.0455
σ_5^2	0.9990	0.0686	1.0074	0.0702

Table 1: Optimal parameters and hyper-parameters and their spread calculated using variational and MCMC approaches.

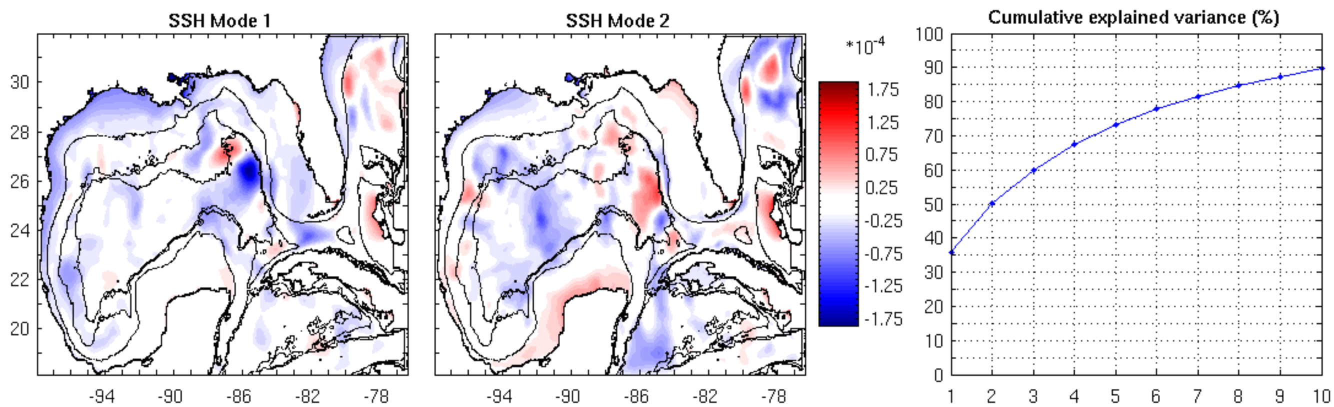

Figure 2: First and Second SSH modes from a 14-day series. The 2 modes account for 50% of variance during these 14 days. The first seems to be dynamically related to the presence of a frontal eddy in the Loop Current region. The second mode exhibits significant energy in the same region but does not seem to lend itself to a simple dynamical interpretation.

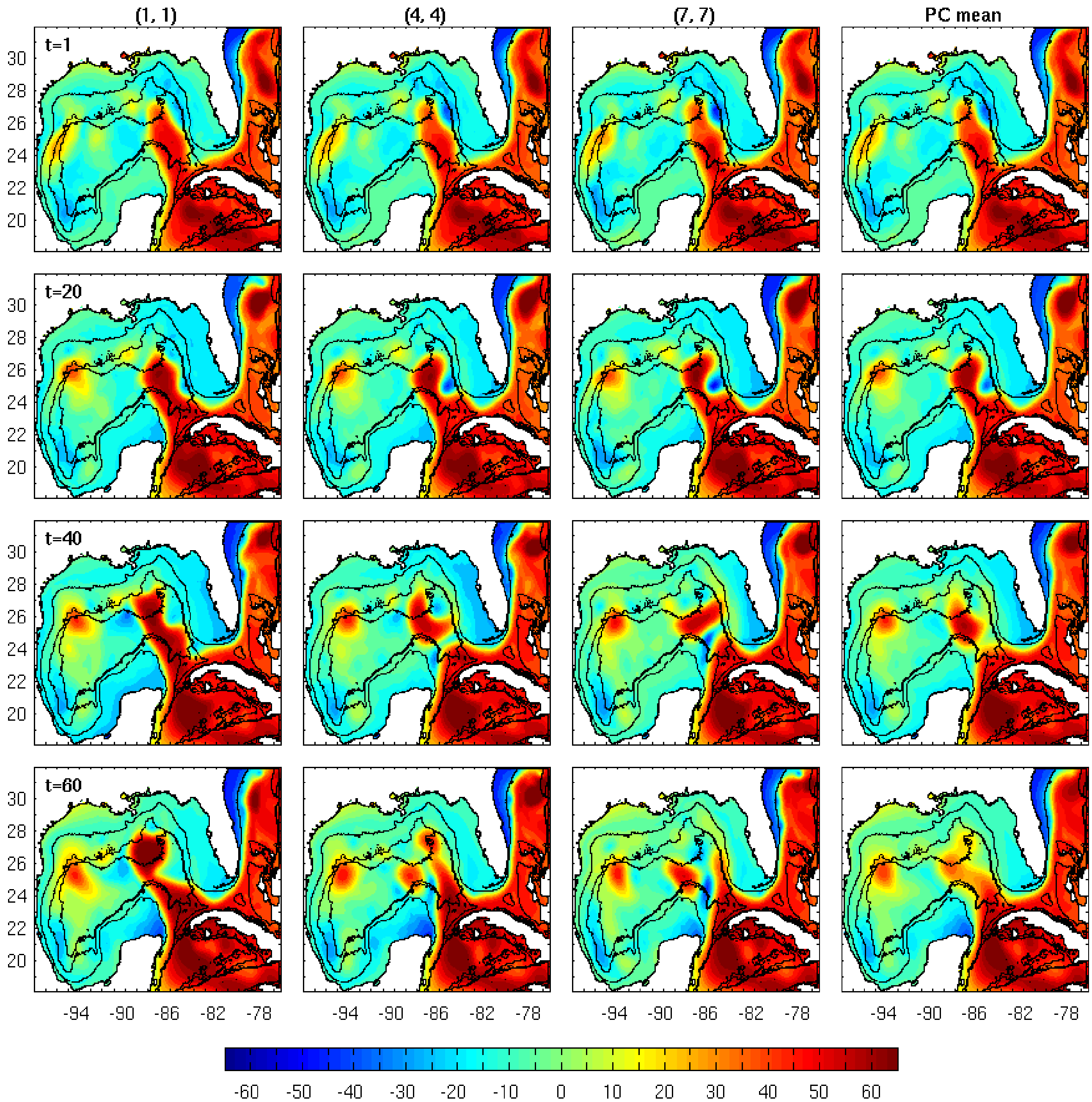

Figure 3: SSH evolution for several HYCOM realizations. The run with the most negative perturbation (first column) exhibiting the weakest frontal eddy, the control run (second column) is associated with a weaker frontal eddy while the third most column has the strongest positive perturbations of the 2 EOF modes and exhibits the strongest frontal eddy. The right most column is the mean SSH extracted from the PC series.

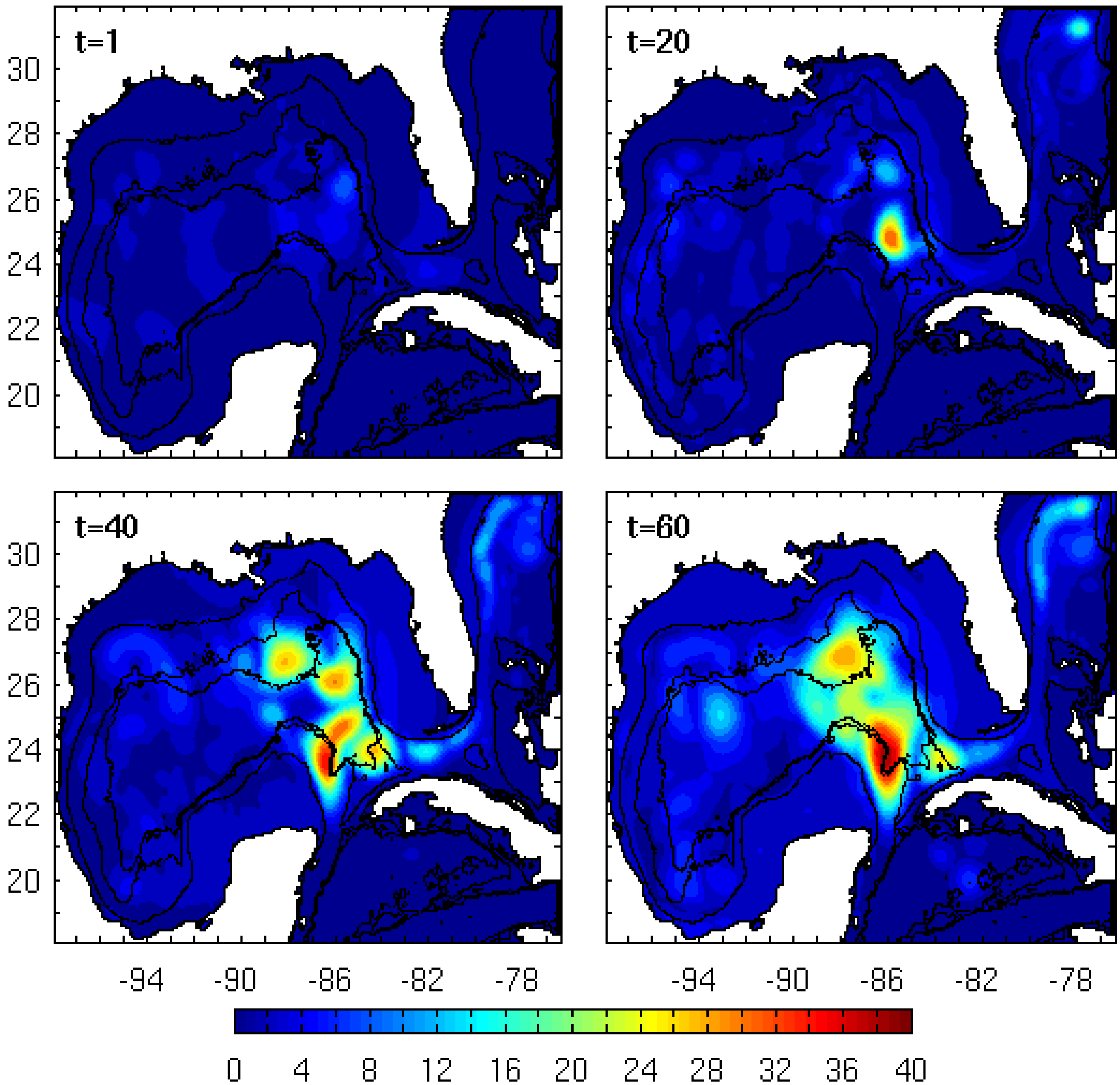

Figure 4: Evolution of the SSH standard deviation's spatial distribution due to uncertainty in initial conditions. The uncertainty maximum at day 20 seems to be related to the early eddy shedding event associated with the strongest frontal eddy realization. Most of the uncertainty is located in the Loop Current region and increases with time.

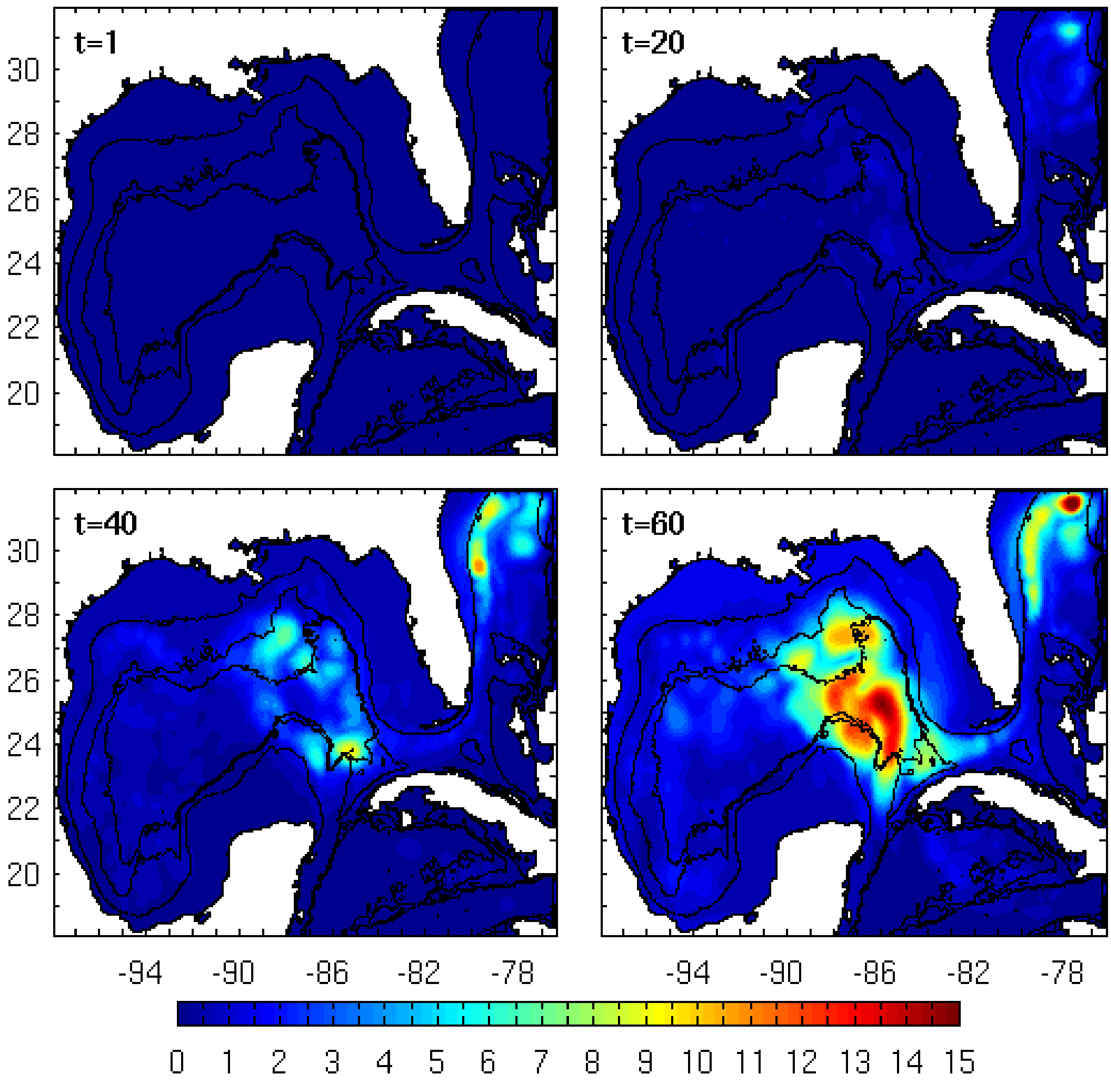

Figure 5: Norm of the PC representation error, $\|\varepsilon\|_2^2 = \sum_q \left[\eta(\vec{x},t,\xi_q) - \eta_{PC}(\vec{x},t,\xi_q)\right]^2 \omega_q$***, for different forecast days. SSH PC-errors (cm) grow in time with maxima in the Loop Current region. On day 60 the PC-error is about 38% of standard deviation and indicates a potential underestimation of the uncertainty.***

IMPACT/APPLICATIONS

The present project presents an approach to characterize the entire response surface of an ocean model to uncertainties in its input data. This has implications for the fields of parameter estimation, and data assimilation, particularly for ensemble Kalman filter based approaches. The methodology developed here will be of use either for the efficient update of the covariance matrices and/or quantifying the errors incurred by small size ensembles. We are currently exploring these ideas.

TRANSITIONS

RELATED PROJECTS

Dr. Ashwanth Srinivasan was partially supported by an NSF-RAPID grant (NSF OCE-1048697) for his work on the oil-fate model during the first year of the project. The results from the present proposal are directly contributing to two grants from the Gulf of Mexico Research Initiative, namely one through the Consortium for Advanced Research on Transport of Hydrocarbon in the Environment (CARTHE) and the Deep Sea to Coast Connectivity in the Eastern Gulf of Mexico (Deep-C) consortium. The uncertainty quantification tools developed here are being applied to oil-fate modelling.

PUBLICATIONS

A. Alexanderian, O. Le Maître, H. Najm, M. Iskandarani, and O. Knio. Multiscale stochastic preconditioners in non-intrusive spectral projection. *Journal of Scientific Computing*, 50(2):306–340, 2011. ISSN 0885-7474. doi: 10.1007/s10915-011-9486-2.

A. Alexanderian, J. Winokur, I. Sraj, A. Srinivasan, M. Iskandarani, W. C. Thacker, and O. M. Knio. Global sensitivity analysis in an ocean general circulation model: a sparse spectral projection approach. *Computational Geosciences*, 16(3):757–778, 2012. ISSN 1420-0597. doi: 10.1007/s10596-012-9286-2.

I. Sraj, M. Iskandarani, A. Srinivasan, W. C. Thacker, J. Winokur, A. Alexanderian, C.-Y. Lee, S. S. Chen, and O. M. Knio. Bayesian inference of drag parameters using AXBT data from typhoon Fanapi. *Monthly Weather Review*, 141(7):2347–2366, 2013. doi: 10.1175/MWR-D-12-00228.1.

A. Srinivasan, J. Helgers, C. B. Paris, M. LeHenaff, H. Kang, V. Kourafalou, M. Iskandarani, W. C. Thacker, J. P. Zysman, N. F. Tsinoremas, and O. M. Knio. Many task computing for modeling the fate of oil discharged from the deep water horizon well blowout. In *Many-Task Computing on Grids and Supercomputers (MTAGS), 2010 IEEE Workshop on*, pages 1–7, November, 2010. IEEE. doi: 10.1109/MTAGS.2010.5699424.

W. C. Thacker. Can hydrographic data provide evidence that the rate of oceanic uptake of anthropogenic co2 is increasing? 8(8), August 2013. doi: 71920.doi:10.1371/journal.pone.0071920.

W. C. Thacker, A. Srinivasan, M. Iskandarani, O. M. Knio, and M. L. Henaff. Propagating boundary uncertainties using polynomial expansions. *Ocean Modelling*, 43–44:52–63, 2012. doi: 10.1016/j.ocemod.2011.11.011.

J. Winokur, P. Conrad, I. Sraj, O. Knio, A. Srinivasan, W. Thacker, Y. Marzouk, and M. Iskandarani. A

priori testing of sparse adaptive polynomial chaos expansions using an ocean general circulation model database. *Computational Goesciences*, 2013. doi: 10.1007/s10596-013-9361-3. in press.

REFERENCES

R. Bleck. An oceanic general circulation model framed in hybrid isopycnic cartesian coordinates. *Ocean Modelling*, 4(2):55–88, 2002.

E. P. Chassignet, L. T. Smith, G. R. Halliwell, and R. Bleck. North atlantic simulations with the hybrid coordinate ocean model (hycom): Impact of the vertical coordinate choice, reference pressure, and thermobaricity. *Journal of Physical Oceanography*, 33(12):2504–2526, DEC 2003. ISSN 0022-3670.

E. P. Chassignet, H. E. Hurlburt, O. M. Smedstad, G. R. Halliwell, P. J. Hogan, A. J. Wallcraft, and R. Bleck. Ocean prediction with the hybrid coordinate ocean model (hycom). In E. Chassignet and J. Verron, editors, *Ocean Weather Forecasting: An Integrated View of Oceanography*, pages 413–426, PO BOX 17, 3300 AA DORDRECHT, NETHERLANDS, 2006. SPRINGER. ISBN 1-4020-3981-6. International Summer School of Oceanography, Agelonde, FRANCE, SEP 20-OCT 01, 2004.

G. R. Halliwell. Evaluation of vertical coordinate and vertical mixing algorithms in the hybrid-coordinate ocean model (hycom). *Ocean Modelling*, 7(3-4):285–322, 2004. ISSN 1463-5003. doi: 10.1016/j.ocemod.2003.10.002.

O. P. Le Maître and O. M. Knio. *Spectral Methods for Uncertainty Quantification*. Springer-Verlag, 2010.

I. Sraj, M. Iskandarani, A. Srinivasan, W. C. Thacker, and O. M. Knio. Drag parameter estimation using gradients and hessian from a polynomial chaos model surrogate. *Monthly Weather Review*, 2013, in review. in review.

HONORS/AWARDS/PRIZES

- 13 invited lectures, of which 2 were plenary.
- O. Knio was named Distinguished Professor, July 1, 2012.
- O. Knio was elected member of the Editorial Board of SIAM/ASA Journal on Uncertainty Quantification.